B.Comp. Dissertation

Augmented Reality: Virtual fitting room using Kinect

By

Lan Ziquan

Department of Computer Science

School of Computing

National University of Singapore

2011/2012

B.Comp.Dissertation

Augmented Reality: Virtual fitting room using Kinect

Project No: H133280

By

Lan Ziquan

U087254R

Supervised by

A/Professor Shengdong ZHAO

Department of Computer Science

School of Computing

National University of Singapore

2011/2012

Deliverable:

Report: 1 Volume

ABSTRACT

This project implements a real-time virtual trying on system using Kinect as input device. Besides the basic feature of trying clothes, the Virtual Fitting Room system also supports features of photo taking and downloading via quick response 2D barcode (QR code). The user interface design evaluation and final system evaluation have been done carefully to show the system usability. The project purpose is to enhance users' shopping experiences by using the system, so that they can spend less time on queuing for fitting rooms and can also easily share their appearances after downloading their photos.

Subject Descriptors:

H.5.1 Multimedia Information Systems

H.5.1 User Interfaces

I.3.7 Three-Dimensional Graphics and Realism

Keywords:

Virtual Try On, Augmented Reality, Computer Graphics, Kinect, Gesture, QR code, User Interface Design

ACKNOWLEDGEMENT

This dissertation could not have been finished without my supervisor, Assistant Professor Shengdong ZHAO, who set up the project with Microsoft and C.K. Tangs and established the connections between the companies and me. He guides and supports me throughout the project. His help is invaluable indeed.

My special thanks also go to Mr. Wai Keong Sam from C.K. Tangs, Mr. Evangelista Kim from Microsoft and my friend Mr. Chow Jianann, who give great supports and valuable suggestions throughout the implementation process.

Finally, I would like to thank all my friends and family who are always encouraging me with their best wishes.

Lan Ziquan

TALBE OF CONTENTS

1. INTRODUCTION AND MOTIVATION

Shopping for clothes is a common daily activity both in-store and online. An in-store shopper usually tries on some selected clothes and examines how well they fit. While an online shopper usually checks how well the clothes fit the models in the online pictures. That is, a clothes shopper will usually make a decision after he/she perceived the visual image of a given garment. Moreover, rather than decide by themselves, some shoppers would like to ask for his/her friends' opinions by sharing those pictures. However, by observing real life shopping experiences, the decision methods as stated above have some constrains and drawbacks that are elaborated as follows.

Firstly, in a physical store, in order to try on some selected clothes a common practice is to queue up and take turns to use the fitting rooms. Due to the limited number of in-store fitting rooms, shoppers usually have to spend most of their shopping time on queuing up (will be even longer during peak hours). Prolonged waiting time will affect customers' patience, which leads to lower customer satisfaction. To shorten the average waiting time, many clothes shops apply a strategy to constrain the maximum number of items that are allowed to bring into the fitting room at one time. However, this strategy is a double-blade sword. Customers with higher purchasing power would like to try more items at one time. In that case, constraining the maximum number of items definitely turns to be a negative factor, which again leads to lower customer satisfaction. Unfortunately, the situation could be worse. Some shoppers will directly give up the idea of trying on after seeing a long queue outside the fitting rooms.

Secondly, shopping for clothes online has more obvious drawbacks. The selling items are inaccessible to the customers, so it is impossible for the customers to physically try on the clothes on themselves until the items are delivered. Customers always make decisions according to the pictures of models online. However, items look good on models does not imply they look good on the buyers as well. Therefore, the inaccessibility of purchasing items at selection time potentially increases the rate of item returns as a result of poor fit.

To conclude what have been observed, both in-store and online shopping have unavoidable limitations which give rise to customers' unsatisfied shopping experiences as well as the retailers' losses of potential sells. In order to deal with the "lose-lose" situation, people are keen on looking for strategies and techniques. Then, an idea called "Virtual Try On" is put forward. Customers would be able to perceive the visual image of how they look like by trying on clothes virtually, so that they can easily filter out some selections without taking turns to use the fitting rooms. Compare with "Physical Try On", "Virtual Try On" takes much less time. Thus, it increases the shopping efficiency for all customers, hence, enhances the shopping experience.

Thirdly, sharing appearances with friends is not always easy. Shopper who goes to physical store alone can hardly share his/her appearance with friends, unless he/she takes a photo and sends it to his/her friends. Also, online shopper would like to send those online photos to his/her friends for suggestions. Combining the idea of photo sharing with "Virtual Try On", the system shall further enhance the shopping experience.

This project report describes an interactive visual computer system called "Virtual Fitting Room" which implemented the concept of "Virtual Try On" with photo sharing feature by using Microsoft's Kinect sensor. The following sections give further details about this project in terms of its objectives, related works, system design, development tools, implementation iterations, evaluations and feedbacks.

2. PROJECT OBJECTIVES

This project was launched together by Microsoft, C.K. Tangs and NUS. Microsoft is in charge of giving hardware support. C.K. Tangs is a sponsor as well as a client. On behalf of NUS, I am the developer whose main objective is to fulfill the client's requirements. The initial requirement is to implement a "Virtual Try On" system using Kinect sensor as its input device. An additional requirement asked during the development process is the feature of photo sharing. This section explains the choice of using Kinect as well as my concerns about the user interface design, the system interactivity and the photo sharing function.

As a motion sensing input device, Kinect is a right tool to be introduced into such a system which needs to track users' positions, poses and gestures. By using Kinect to interact with the system, the user does not need to use any additional control devices. Therefore, interaction between the user and the system becomes more natural and convenient.

Virtual Fitting Room is an interactive system that mimics the real fitting experiences. Displaying user's whole body image is a practical concern. Unlike a usual user interface, the screen requires a large portion of center space to display the user's mirror image. Then, the remaining part of screen should be utilized effectively in order to display other information, such as instructions and functional menus. Detailed interface design is described in the system design section.

Besides the screen management issue, system interactivity needs to be concerned carefully as well. To be more specific, this system needs to focus on recognizing gestures in a reliable manner. A reliable gesture recognizing method should minimize both false positive rate and false negative rate. For example, certain gesture is predefined as a selection gesture. When user intends to select an item and performs the predefined selection gesture, system should be able to detect the gesture and give its response accordingly. However, sometimes user performs unintended selection gesture that is also detected by the system, so the gesture recognizing method should be designed carefully in order to reduce the happening of accidental selections.

In addition to the two primal aspects, i.e. user interface design and system interactivity, photo sharing function is the highlight of the whole system. Unlike usual photo taking scenario, the system can help the user to take a photo (by screenshot) and send the photo to the user. Since the system is going to be deployed in a public place, it is important to use an easy way to enable an arbitrary customer to retrieve his/her photo. Asking for customer's personal information, such as email address or cell-phone number, is troublesome and inefficient. Instead of pushing the photo to the customer's personal device (passive retrieval), the system may allow the customer to pull/download the photo in some quick way (initiative retrieval). Using Quick Response 2D-barcode (QR code) is my solution. Nowadays, QR code is widely used. Many free "QR scanner" apps are available in the app stores. One can use a smart phone to scan the QR code which encodes a link to the photo in order to download the photo easily and quickly.

To sum up, the project objectives are divided into several sub-tasks: study of Kinect's functionalities, user interface design, gesture definition, gesture recognizing, photo taking (screen shot), photo uploading and QR code generation. Section 5 describes the study of implementation for each sub-task in detail.

3. RELATED WORKS

Since this project touches many researching areas, this section divides the related works into four categories: virtual trying, computer visualization, gesture control and decision making for clothes selection.

3.1 VIRTUAL TRYING

Many virtual trying systems use the 3D avatar approach. Some researches work on making garment design process easier by manipulating virtual clothes (Volino, Cordier, & Magnenat-Thalmann, 2005; Wacker, Keckeisen, Stoev, & Straßer, 2003). Some other researches focus on making virtual humans and virtual clothes look more real (Fuhrmann, Groß, Luchas, & Andreas, 2003; Meng, Mok, & Jin, 2010). Internet based garment design is also interested by the researches. That gives chance for online shoppers to dress their avatar while doing online garment purchasing (Protopsaltou, Luible, Arevalo, & Magnenat-Thalmann, 2002; Cho, Park, Boeing, & Hingston, 2010). Moreover, some projects also find ways for the system to automatically generate 3D human avatars from a set of static images (Magnenat-Thalmann, Seo, & Cordier, 2004).

Without using avatars, people come up with several different approaches to change clothes in real-time. Some systems make use of markers on the shirt to change clothes patterns by overlying different patterns on the mirror images (Ehara & Saito, 2005), while some other systems change the original colors of clothes or shoes to make them look different (Cheng, et al., 2008; Eisert, Rurainsky, & Fechteler, 2007).

3.2 COMPUTER VISION

Before the invention of Kinect, people used one or multiple video cameras in computer vision systems (Poppe, 2007). Using computer vision to track human body motions has been studied from early years until today (Wren, Azarbayejani, Trevor, & Pentland, 1997; Maes, Darrel, Blumberg, & Pentland, 1995). Researchers come up with various approaches and algorithms to help the computer system recognize human bodies (Lee & Nevatia, 2007; Lin, Wnag, Liu, Xiong, & Zeng, 2009). Some systems use more than one video camera for accurate modeling (Gavrila & Davis, 1996), while some other systems introduce depth value into the system as what Kinect does (Du, et al., 2011). Recently, in the year 2011, there are also a few papers that make use of Kinect for augmented reality project (Vera, Gimeno, Coma, & Fernández, 2011) or investigate the accuracy of Kinect depth data (Khoshelham, 2011). I believe more and more researches will work with Kinect in the future in the area of computer vision.

3.3 GESTURE CONTROL

Many studies focus on recognizing static multi-finger hand gestures (Pavlovic, Sharma, & Huang, 1997; Malik & Laszlo, 2004) or a sequence of human body gestures (Kim & Kim, 2006), while there are also some other studies that propose accelerometer-based gesture control (Kallio, Kela, Jani, & Plomp, 2006). Besides, some projects also concern using gestures to control virtual human structures in the virtual space as this project does (Huang & Kallmann, 2009). A main purpose of this project is to design a gesture that is easily learned by most of the people, possibility including elderly and disabled (Bhuiyan & Picking, Gesture Controlled User Interfaces for elderly and Disabled, 2009; Bhuiyan & Picking, Gesture Controlled User Interfaces for Inclusive Design, 2009). An easy and intuitive gesture can help to increase the usability of the system (Bhuiyan & Picking, Gesture Control User Interface, what have we done and what's next?, 2009; Bhuiyan & Picking, A Gesture Controlled User Interface for Inclusive Design and Evaluative Study of Its Usability, 2011).

3.4 DECISION MAKING FOR CLOTHES SELECTION

Different types of researches explore the area of helping making decisions. Some system designs make the comparing process more convenient for the user by showing a set of similar existing images (Zhang, Matsumoto, Liu, Chu, & Begole, 2008; Chu, Dalal, Walendowski, & Begole, 2010). Some systems use artificial intelligence technology to give recommendations according to the history results or by communicating with the user (Santangelo, et al., 2007; Shen, Lieberman, & Lam, 2007). Social factor is also introduced in some researches. Those projects combine the systems with the social networks (Tsujita, Tsukaka, Kambara, & Siio, 2010; Liew, Kaziunas, Liu, & Zhuo, 2011).

4. SYSTEM DESIGN

4.1 ABSTRACT SYSTEM WORK FLOW

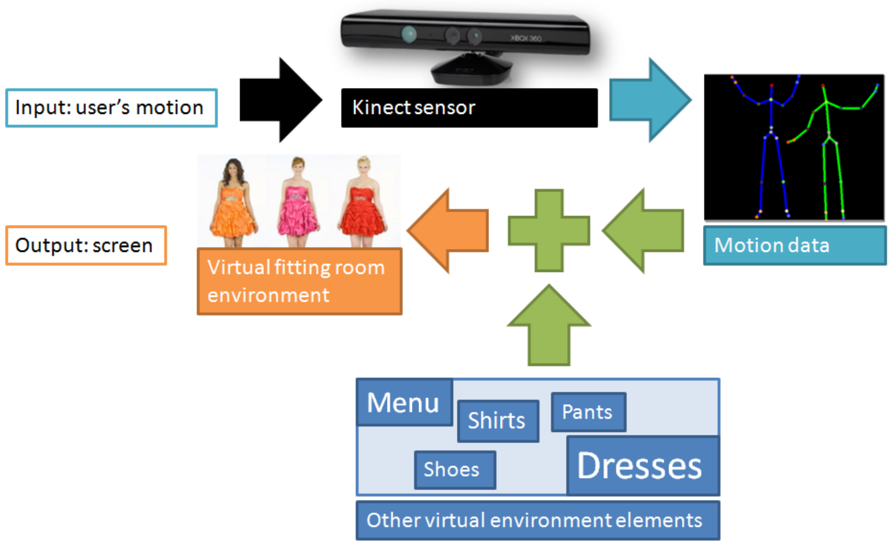

Figure 1. Abstract System Work Flow

Figure 1 shows an abstract system work flow of the "Virtual Fitting Room". The system takes the user's motion as input, and converts the user's motion to motion data via Kinect sensor. Together with the input motion data, other virtual environment elements, like user interface components and dress models, are integrated as a virtual scene. Finally, the integrated virtual scene is rendered by the graphic card and is displayed on the screen. This section shows the system preparation and the software design. The system preparation consists of several components: hardware setup, programming environment, programming language and API specifications. The software design part shows detailed user interface designs and system underlying work flow in terms of its sequence diagram.

4.2 HARDWARE SPECIFICATION

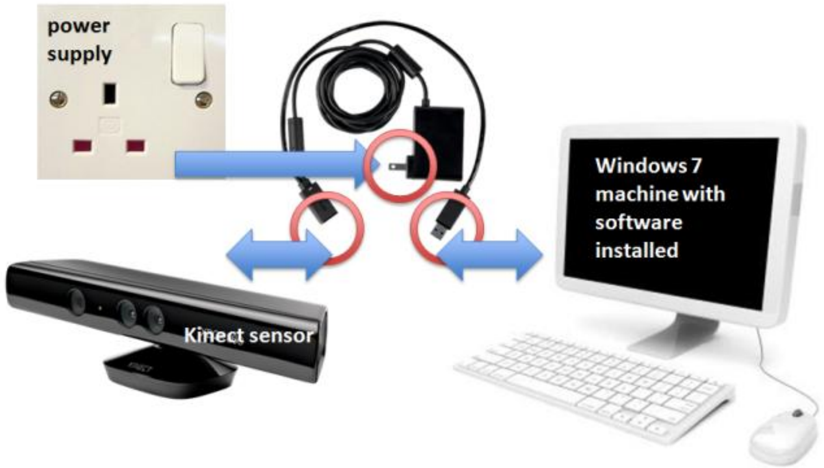

Figure 2. System Hardware Setup

Figure 2 shows the structure of the system hardware setup which is very simple. It only requires a Kinect sensor, a PC machine and a Kinect external power line. The external power line not only provides sufficient power to the Kinect sensor but also extends Kinect sensor's data cable with a USB cable in order to connect with the PC machine.

4.3 ENVIRONMENT AND IDE SPECIFICATION

KinectSDK (beta) is only compatible with Windows 7 operating system, Visual Studio 2010 and .Net Framework 4.0. Therefore, I chose to use Windows 7 (x86), Visual Studio 2010 (Professional Edition) for the software development

4.4 PROGRAMMING LANGUAGE AND API SPECIFICATIONS

KinectSDK provides libraries for both C++ and C#, whereas, the drawing libraries, such as OpenGL and DirectX, were not compatible with Visual Studio 2010 by using C#. Therefore, I chose to use C++ as the programming language and DirectX9.0 as the API for rendering 3D virtual scenes. In addition, compare with Java and C#, C++ does not have built-in FTP data transferring APIs, so I found a useful library called "Curl" for easy data transferring using C++. Section 5 describes how these libraries are being used in this system.

4.5 USER INTERFACE DESIGN

Figure 3, 4, 5 are the three versions of user interface designs during the iterative implementation process. The interface evolves according to the interface evaluation results as well as the client's requirements. The set of main functionalities still remains the same, which are selecting clothes, changing to a different set of clothes displayed and taking photo.

Figure 3. User interface v.1

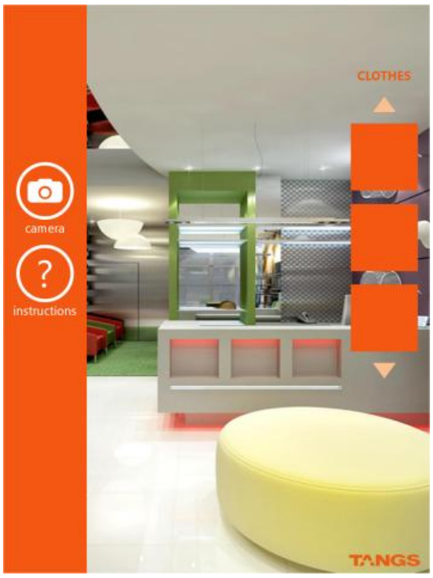

Figure 4. User interface v.2

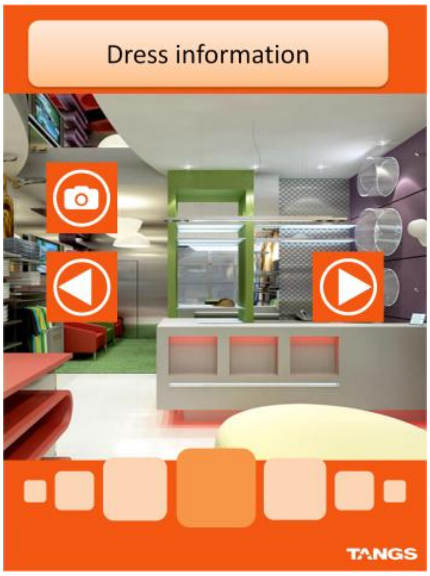

Figure 5. User interface v.3

In version 1, the screen layout is designed to be landscape. As stated in section 2, a large portion of center screen space is used to show the mirror image of the user, and then the two sides are used to display the functional components. The functional components are divided into clothes selection part and photo button part. Compare with taking photo, selecting clothes is a more frequent activity. Since there are more right-handed people than left-handed ones in Singapore, clothes selection part is placed on the right hand side while photo button is placed on the left hand side.

Version 2 changes the screen orientation from landscape to portrait. This design makes the interface look more like a mirror. The interface color tone becomes orange so that it matches the style of Tangs. The clothes selection part is still on the right hand side and the functional buttons are on the left hand side.

Version 3 is the most recent interface design, which looks quite different from the previous two. The screen orientation remains portrait and the color tone remains orange, while the mirror image area changes from a rectangle shape to a square one. This new design makes more use of the top and bottom parts of the screen: the top part is used to display some system instructions and the clothes information; the bottom part is allocated by a horizontal list of clothes snap shots which were previously displayed on the right hand side. In the previous versions, the user changes his/her current trying by explicitly selecting a new dress snap shot from the right hand side list. In the version, the user changes his/her current trying automatically by selecting the left and right arrow buttons. The centered one displayed in the clothes list is the current dress that the user is trying on.

The interface evaluation for version 1 is presented in section 7. The reasons of keeping modifying interface designs are explained in detail.

4.6 UNDERLYING WORK FLOW

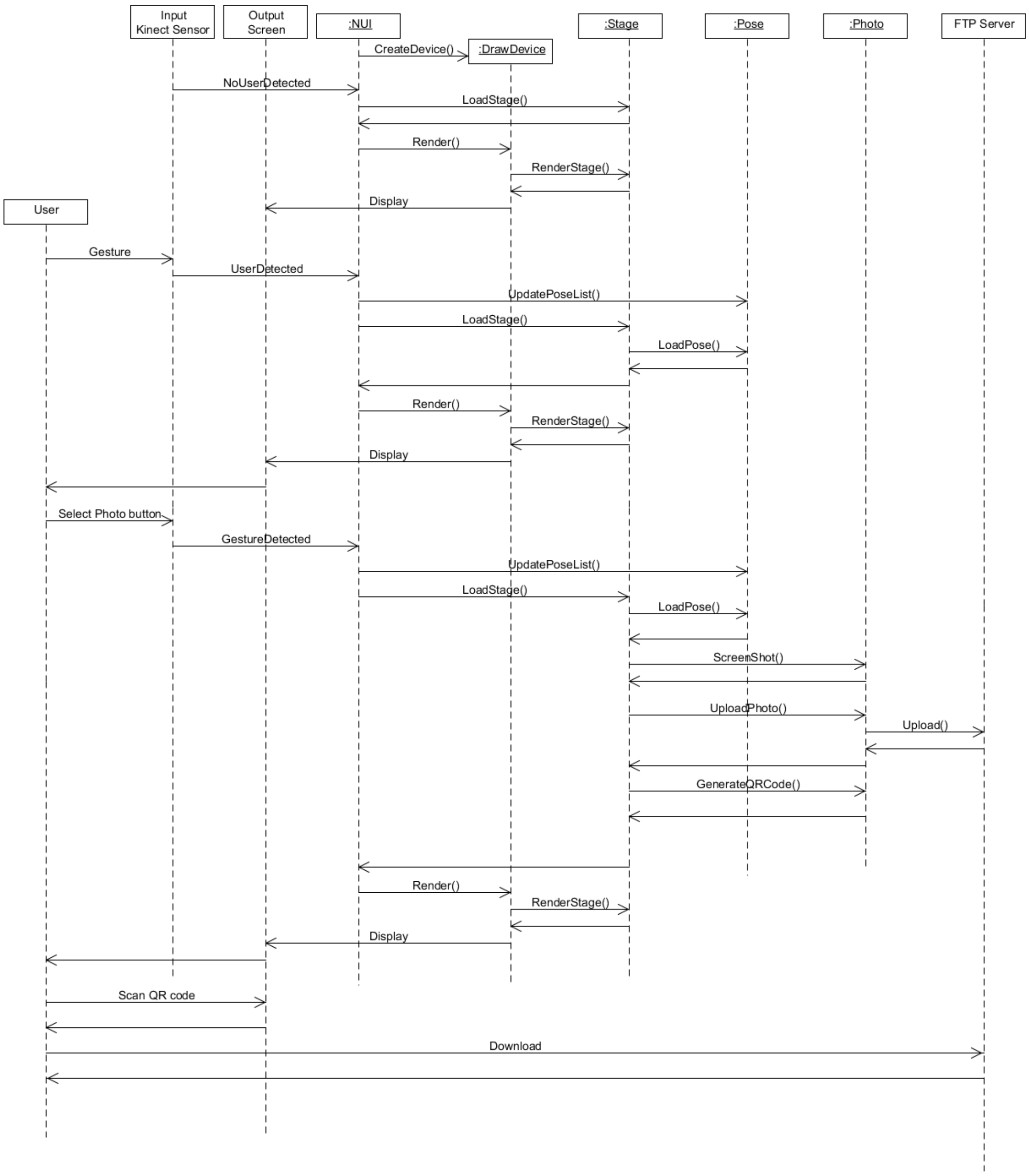

Figure 6. Simplified System Sequence Diagram

Figure 6 is a simplified version of the system sequence diagram. Except User, Kinect sensor, Screen and the FTP server, the system contains four main components: NUI, DrawDevice, Stage and Pose.

NUI is the component that communicates with Kinect Sensor. Depends on whether a user is detected, NUI asks Stage components to prepare different scenes and asks DrawDevice to render the prepared scenes.

DrawDevice is the component that performs displaying tasks using the graphic card. It rendered whatever 3D virtual scenes that were prepared by Stage onto the screen.

Stage is the component that manages what should be displayed. If there is no user detected, Stage does not need to prepare any user motion data into current scene. Otherwise, Stage loads current user position data and integrates the data with virtual clothes to form the current scene. Meanwhile, Stage also asks Pose to detect gestures and updates Stage status accordingly.

Pose is the component that stores and analyzes motion data. Pose keeps updating motion data history once someone is detected. It also contains functions to recognize gestures.

Photo is the component that takes charge of everything about photo taking. It takes screen shot and uploads the image onto an FTP server. Also, it generates an image of

QR code that encodes the link to the uploaded image. Hence, Stage can use the QR code image for its preparation.

To conclude, this section presents different levels of system designs: abstract system work flow, system setup preparation, user interface design and system underlying work flow.

5. STUDY OF DEVELOPMENT LIBRARIES AND PACKAGES

5.1 KINECTSDK(BETA)

In order to implement the system, the first step is to study the Kinect's functionalities. However, because Kinect is new, there is very few of references of Kinect projects. Fortunately, Microsoft has released a non-commercial KinectSDK (beta) on June 16, 2011. With the help of KinectSDK programming guide, the basic functionalities of Kinect sensor is learned and summarized as follows.

Kinect sensor generates three sets of data streams from its working environment: color image stream, depth data stream and skeleton data.

5.1.1 COLOR STREAM

Each frame of color stream contains an image that is similar to a normal video camera image. Figure 7 is a sample image provided in the programming guide.

Figure 7. Kinect color data stream

5.1.2 DEPTH STREAM

Each frame of depth data stream gives a set of coordinates which contains not only X and Y direction values but also Z direction (depth) values. Figure 8 is a sample image generated from a frame of depth data stream. Two different players are indentified and distinguished by two different colors. Figure 9 shows the coordinate space given by Kinect.

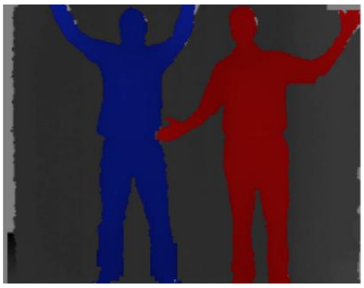

Figure 8. Kinect depth data visualization

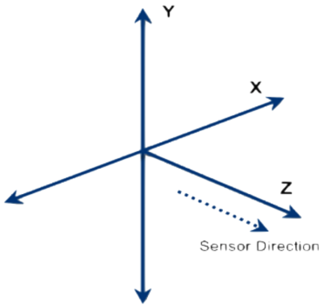

Figure 9. Kinect sensor direction

5.1.3 SKELETON DATA

Each frame of skeleton data is converted from a frame of depth data stream by KinectSDK library. A frame of skeleton data contains a set of 3D coordinates of user body joints defined in the way as shown in Figure 11. Figure 10 is a sample image that is drawn by using a frame of skeleton data.

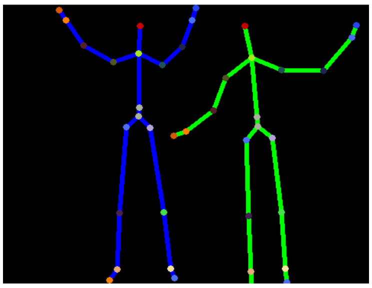

Figure 10. Kinect skeleton data visualization

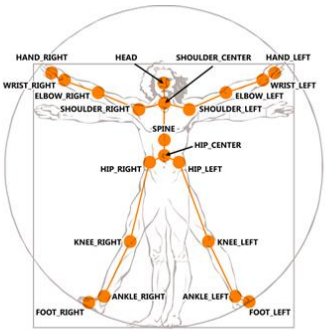

Figure 11. Joints definition

5.1.4 OTHER KENECT SENSOR SPECIFICATIONS

Besides the three kinds of input data streams, some other sensor specifications are also important to know.

According to the programming guide, Kinect's skeletal tracking distance is between 1.2 to 3.5 meters. After mounted with a Kinect Zoom produced by Nyko, the tracking distance is reduced up to 40% closer, i.e. about 0.7 meter.

Kinect viewing angle is given as a 43 degrees vertical by 57 degrees horizontal field of view. This information helps to specify the parameters of the 3D viewing frustum while rendering the virtual scene.

The best resolution given by the color stream is 1024 by 768 which is not good enough for displaying on a large screen. However, the frame rate is about 30Hz which is sufficient to support a smooth video display.

5.2 DIRECTX9.0

Virtual Fitting Room requires 3D rendering technique to simulate the virtual environment. As a popular 3D rendering API, DirectX is a proper library to be studied and used in the project.

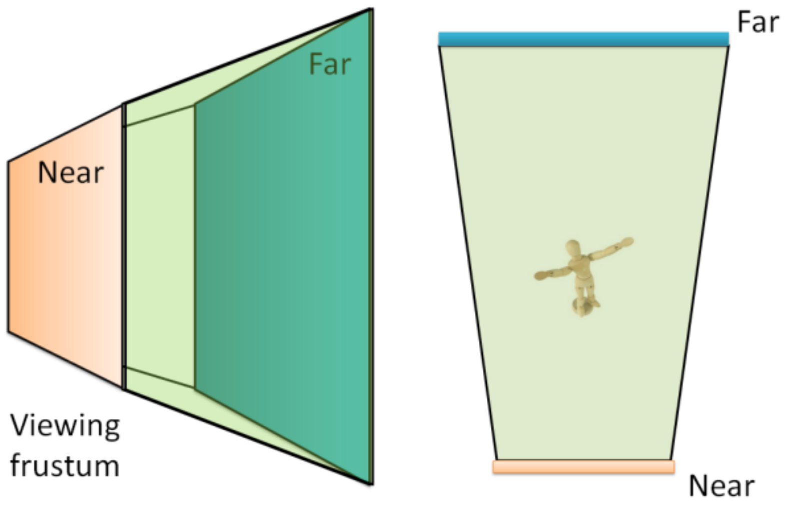

Figure 12. System viewing frustum design

Figure 12 shows the basic structure of the 3D virtual scene. The far plane is used to display the color video frame image. The near plane is used to draw the user interface. The models of clothes are rendered in the space between the far and near planes. With this arrangement, the final output looks the same as the expected design: the user interface components are displayed to cover everything else; the clothes models are displayed to cover on the human body which is displayed in the color video frame image.

Creating clothes models is a big deal, because it requires a good sense of art as well as the knowledge of clothes animation which is another researching area. Hence, in order to fulfill the project objectives first, I chose to use 2D clothes textures instead of 3D clothes models. The 2D clothes textures need to be prepared well, so that the clothes would fit the users' bodies nicely. Figure 13 shows the process of clothes textures preparation: the first step is to remove the background from the original photo; the second step is to fit the dress on a fixed 2D mannequin; the final step is to save the output texture. After this preparation process, all the clothes textures can be rendered with the same set of parameters in the 3D virtual scene.

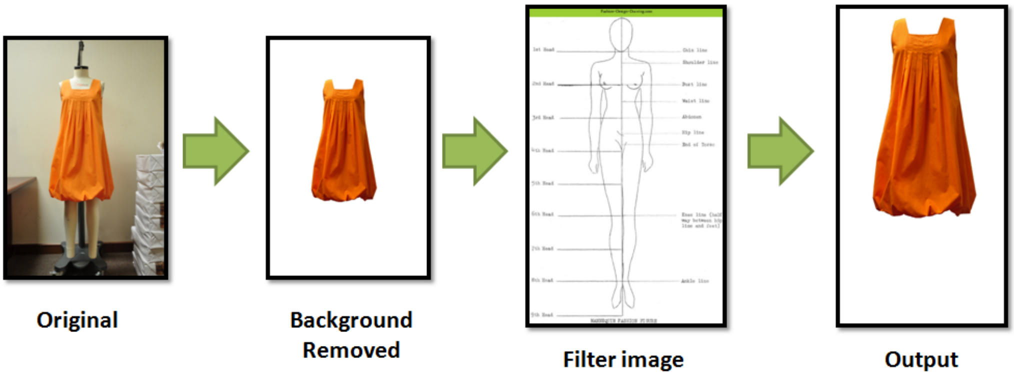

Figure 13. Texture preparation process

Rendering 2D clothes textures in the virtual scene requires at least one reference point's coordinate values and two reference vectors. The reference point is used to determine where to put the clothes texture and it is chosen to be the shoulder center joint defined by KinectSDK. A left-right reference vector and a up-down reference vector are used to determine the orientation of the clothes texture: the left-right reference vector is chosen to be the one from the right shoulder joint to the left shoulder joint (as shown in Figure 14 in green); and the up-down reference vector is chosen to be the one from the shoulder center joint to the spine joint (as shown in Figure 14 in purple).

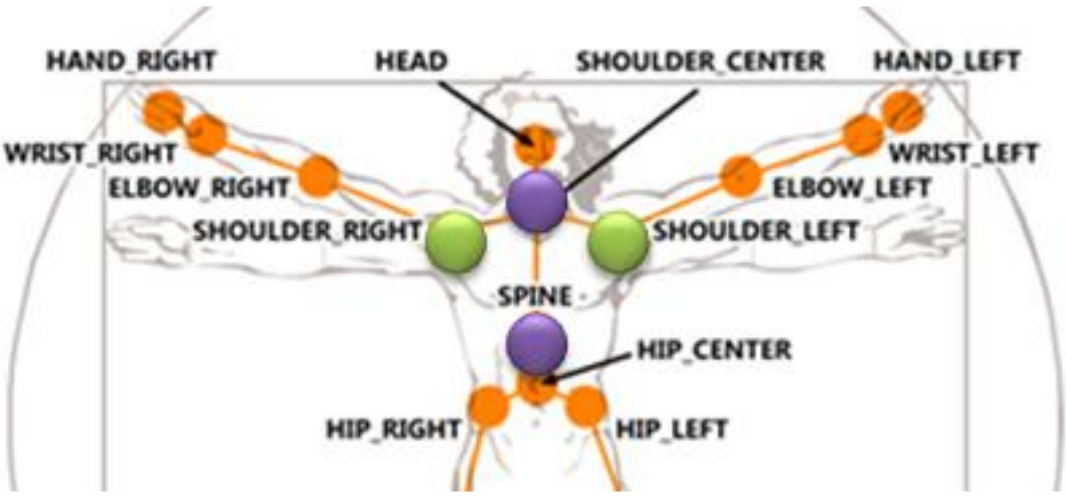

Figure 14. Points of reference vectors on body joints set

5.3 QR CODE GENERATION AND CURL LIBRARY

There are a lot of ways of generating QR codes. Among them, I discovered the easiest one provided by qrserver.com which is getting the QR code image from the following link: http://api.qrserver.com/v1/create-qr-code/?data=<message>&size=400x400

Since C++ does not have any built-in libraries for FTP data transfers, I found the Curl library for uploading files onto a FTP server.

6. IMPLEMENTATION

Virtual Fitting Room system is developed in an iterative approach. This section records the implementation iterations one by one.

6.1 ITERATION 1: RENDERING COMPONENTS TOGETHER

As described in section 5, the virtual scene consists of three components: user interface component, clothes model component and color frame image component. Figure 15 is a screen shot taken at the first time when the three components are rendered together. It is clear to see from Figure 15 that the color frame image is display as the background of the scene; there colored rectangles are displayed as interface buttons; a 3D model constructed with yellow triangles covers the user in the color frame image.

The light was dimmed on purpose in order to understand that generating skeleton data only depends on the depth stream data instead of depending on the combination of color stream and depth stream.

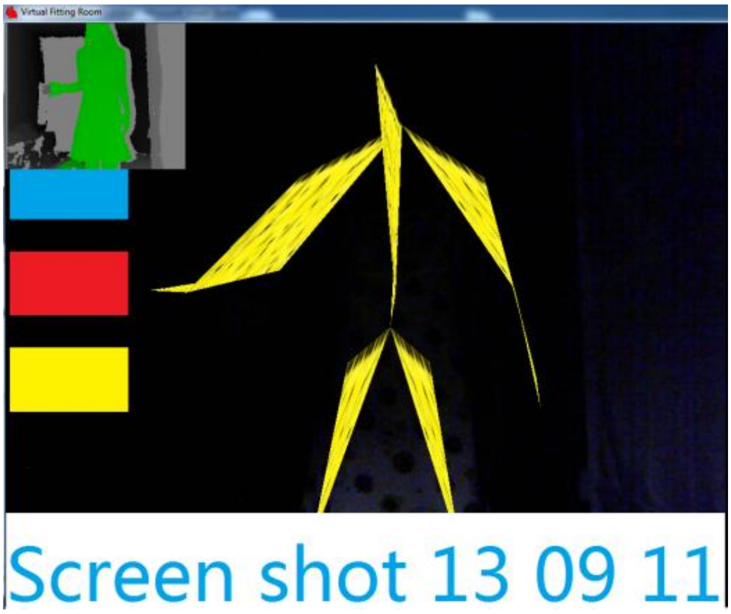

Figure 15. Screen shot in 13.09.2011

Figure 16. Screen shot in 04.10.2011

6.2 ITERATION 2: THE FAILURE OF CREATING 3D CLOTHES MODELS

At first, I wished to create 3D clothes models whose shapes can be changed dynamically according to the change of skeleton data. After spending several weeks on doing it, I was still not able to create a nice clothes model. Therefore, I chose to an alternative approach by using 2D clothes textures. Figure 16 is a screen shot taken at the point of time when I stopped developing the 3D clothes models.

6.3 ITERATION 3: COMING UP WITH THE 2D CLOTHES TEXTURE PREPARATION PROCESS

The preparation process is described in section 5. From this iteration onward, this process is used to prepare clothes textures for all the following iterations.

6.4 ITERATION 4: ENABLING GESTURE AND IMPROVING USER INTERFACE

In this iteration, the system is able to detect simple selection gesture which is defined to be hovering right hand on a button for 1 second. A hand shaped cursor is introduced to indicate the position of the right hand on the menu. The user is allowed to try on some clothes by selecting the corresponding snap shots. Figure 17 is a screen shot that shows the improved interface as well as how the 2D clothes texture looks like on the user.

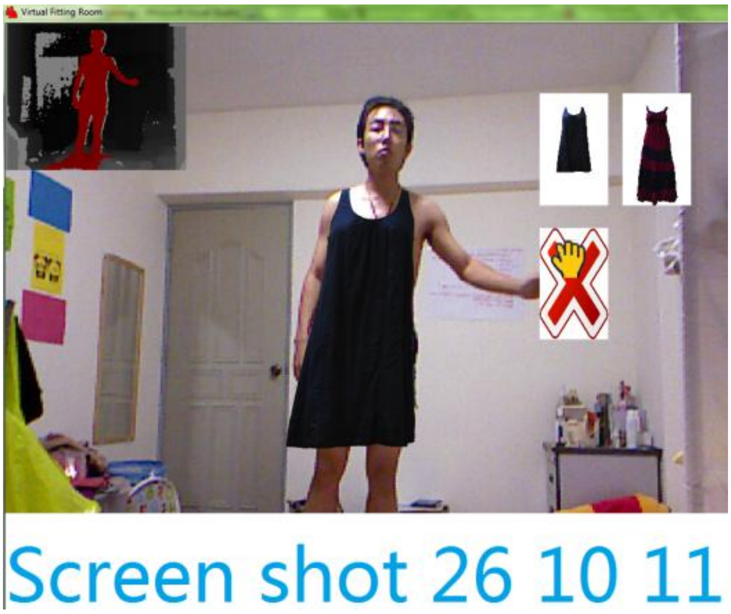

Figure 17. Screen shot in 26.10.2011

6.5 ITERATION 5: ADDING PHOTO TAKING FUNCTION AND IMPROVING USER INTERFACE

Photo taking function is added in this iteration. Figure 18 is a screen shot taken by selecting the photo taking button while interacting with the system. The interface looks similar to the design of version 1, whereas the photo taking button is on the top instead of on the left hand side because only right hand selection gesture is enabled at that time.

Figure 18. Screen shot in 30.10.2011

This screen shot is taken in C.K. Tangs when I demonstrated the system to the client.

6.6 ITERATION 6: UPLOADING PHOTO, GENERATING QR CODE AND IMPROVING USER INTERFACE

In this iteration, the system is able to upload the screen shot on an FTP server and then generate a QR code from the image URL. The interface version 2 is designed and used in this iteration. Figure 19 is an image produced by the photo taking function. The interface component is removed before taking the screen shot, so that only the color frame image and the selected clothes texture is shown in the photo.

Figure 19. A photo taken by the system

6.7 ITERATION 7: SYSTEM EVALUATION AND IMPROVING USER INTERFACE

Up to iteration 6, the basic features that fulfill the project objectives are all implemented and integrated. This iteration focuses on the system evaluation to explore the system usability as well as the problem of system deployment. As a result of this iteration, the interface version 3 is designed and being developed. Section 7 presents the system evaluation process and also gives the reasons of interface modifications.

7. EVALUATION AND FEEDBACK

Two types of user evaluations have been done during the process of iterative implementation: one is the interface conceptual model evaluation; and the other is the system evaluation.

7.1 INTERFACE CONCEPTUAL MODEL EVALUATION

The purpose of the interface evaluation is to come up with a good user interface before the basic features are implemented. Five evaluators (2 females and 3 males) are invited to try and play with a mock up interface in this evaluation. Each evaluator is asked to fill a questionnaire after using the interface conceptual model. The questionnaire asks some system questions such as screen management and learning factor, and also asks some general questions such as the number of functionalities tried and the number of clothes tried. The evaluation result is shown as follows.

- Overall, 2 out of 5 are not satisfied with the interface design at all.
- 4 out of 5 are clear about the screen layout and the rest one is confused.
- All of the 5 evaluators think it is easy or very easy to use the interface.
- 4 out of 5 try all the functionalities from the interface and the rest one does not try the QR code only.
- All of the 5 evaluators try all the buttons to see different images and their reasons are quite similar, i.e. curiosity.
- 3 out of 5 evaluators point out that the interface is lack of professionalism.

- One evaluator points out that the list of buttons can be designed in a cyclic manner so that it is more flexible to navigate.

Besides the questionnaire evaluation, the interface version 1 (implemented during iteration 5 as shown in Figure 18) is also evaluated by the clients. Some more serious problems are coming out. Because the screen layout is landscape, the down arrow button is too low to be reached by hovering hand on that button.

To summarize the feedbacks, interface version 2 is designed. As described in section 4, version 2 changes the color tone to be orange so that the interface looks no longer naïve. It also changes to portrait orientation so that the all the buttons are easily reachable by hands.

7.2 SYSTEM EVALUATION

The purpose of the system evaluation is to investigate the usability of Virtual Fitting Room. Six evaluators (4 females and 2 males) are invited to try the fully functional system. There are three steps to finish the system evaluation: a before-scenario questionnaire, a list of task walkthrough and an after-scenario questionnaire.

7.2.1 BEFORE-SCENARIO QUESTIONNAIRE

The before-scenario questionnaire asks many questions about shopping habits and two questions about the popularity of QR code. The survey result is listed as follows.

- 5 out of 6 go shopping at least once per week and the rest one goes shopping once per month.
- All of them go shopping on weekend or public holidays or Friday afternoon or holiday eve.
- 3 out of 6 try less than 5 clothes on a shopping day and the rest of them try 5 to 10 clothes.
- 4 out of 6 spend 100 to 500 dollars on clothes and the rest two spend less than 100.
- Interestingly, all the four female evaluators uniformly choose the following three points to make purchasing decisions:
  - Try the clothes on and decide by myself
  - Ask friends who go shopping together
  - Take a photo and show/send it to someone for opinion
- All of them think that queuing for the fitting rooms is one of the most time consuming activities on a shopping day.
- 4 out of 6 think that queuing for payment is also time consuming, and the rest two point out the trips to the shopping malls take a lot of time.
- Only one of them does not know about QR code, and all the other five people have QR scanners in their cell phone.

The result shown above goes over my expectation. Most of the evaluators go shopping frequently and they all feel that queuing is one of the most time consuming activities on a shopping day. The female shoppers like to take photos and share with friends for opinion. Most of them know about QR code and also own QR scanner apps.

7.2.2 SCENARIO TASK WALKTHROUGH

In the next step, I introduce the Virtual Fitting Room system to them and ask them to complete a list of seven simple scenario tasks by using the system.

1. Let the system recognize you so that you can control and play with it.
2. Try on some dresses.
3. Try more dresses by using the navigation arrows.
4. Take a photo.
5. Download and save your photo using your cell phone.
6. Send your photo to a friend.
7. Give up the system control by just walking away.

In this scenario design, each evaluator is asked to go through every basic feature of the system. All of them are asked to interact with the system for at least 10 minutes.

7.2.3 AFTER-SCENARIO QUESTIONNAIRE

After the scenario task walkthrough, six evaluators are asked to fill the other questionnaire to evaluate the usability and satisfaction of the Virtual Fitting Room system. Figure 20, 21, 22, 23, 24 are the distribution results that measure the system usefulness, ease of use, ease of learning, satisfaction and overall scenario evaluation respectively.

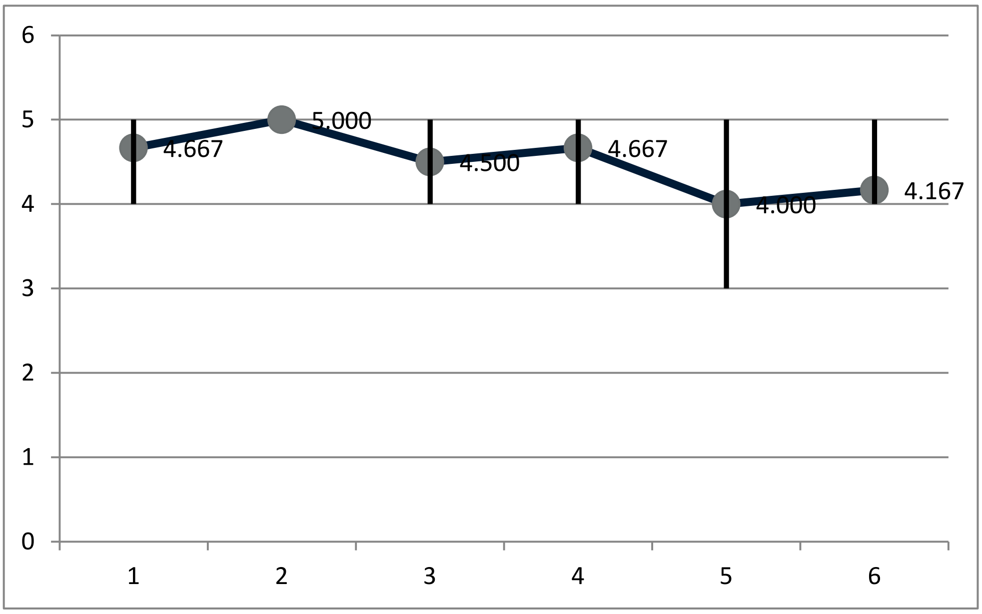

Figure 20. Usefulness

Figure 20 measures the system usefulness. The average score of each question reaches 4.0 or above, which shows that the Virtual Fitting Room system is useful and meets general users’ needs.

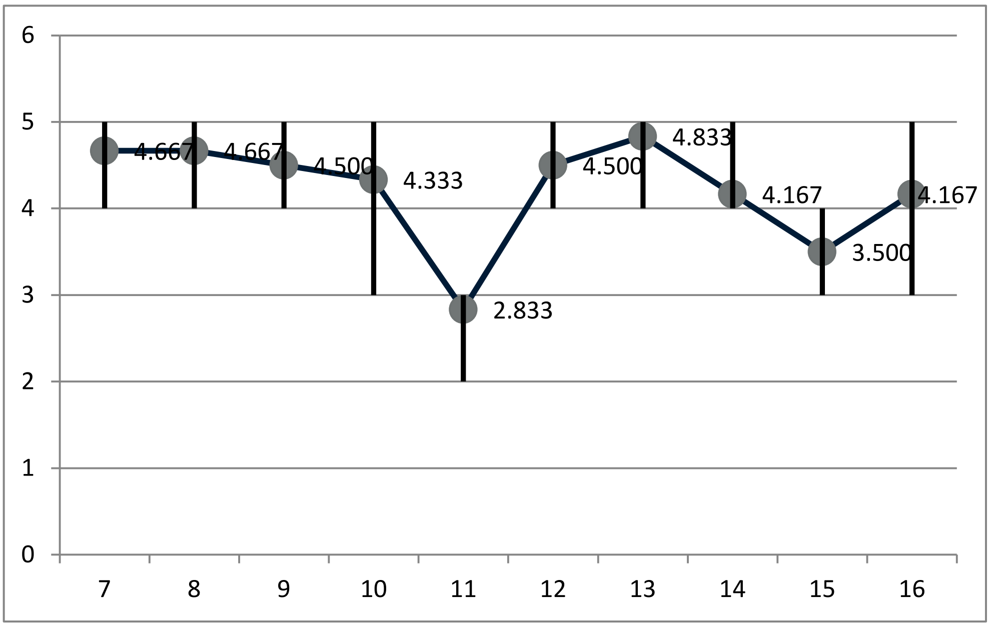

Figure 21. Ease of Use

In Figure 21, except question 11 and question 15, the average scores of the other questions are above 4.0, which indicate the system is relatively easy to use. Question 11 asks whether the system is flexible. A score of 2.833 out of 5 shows that the system is not very flexible. A female evaluator who gives a score of 2 to the question 11 tells that is mainly because the clothes textures are flat and the system does not allow her to turn around or see her back. Question 15 is asking if both occasional and regular users would like the system. The same evaluator who gives the lowest score to question 11 also gives the lowest score of 3 to this question, and her reason is more or less the same as for question 11.

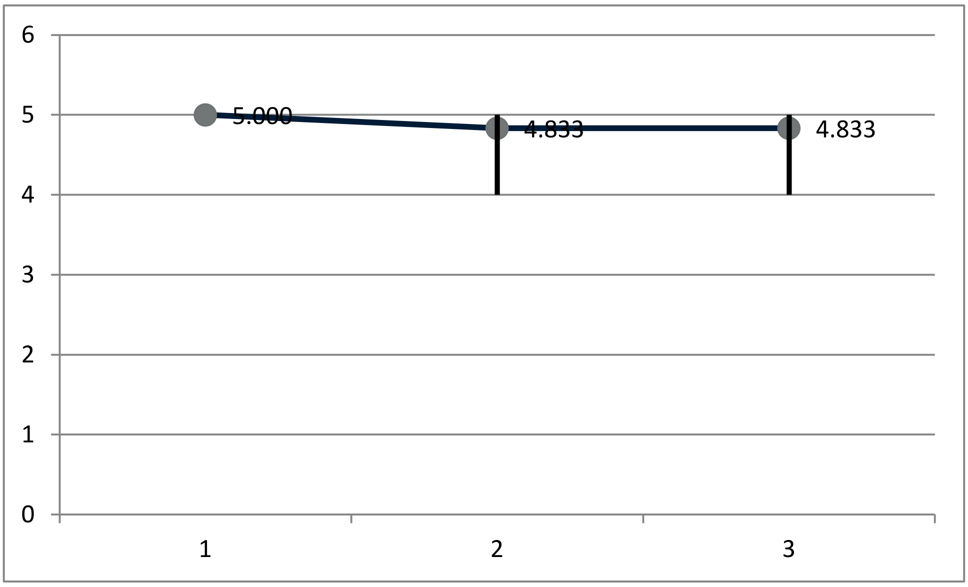

Figure 22. Ease of Learning

Figure 22 shows the measurement of ease of learning. Because the system features is not complicated and the gesture used is intuitive and straightforward, the system is very easy to learn.

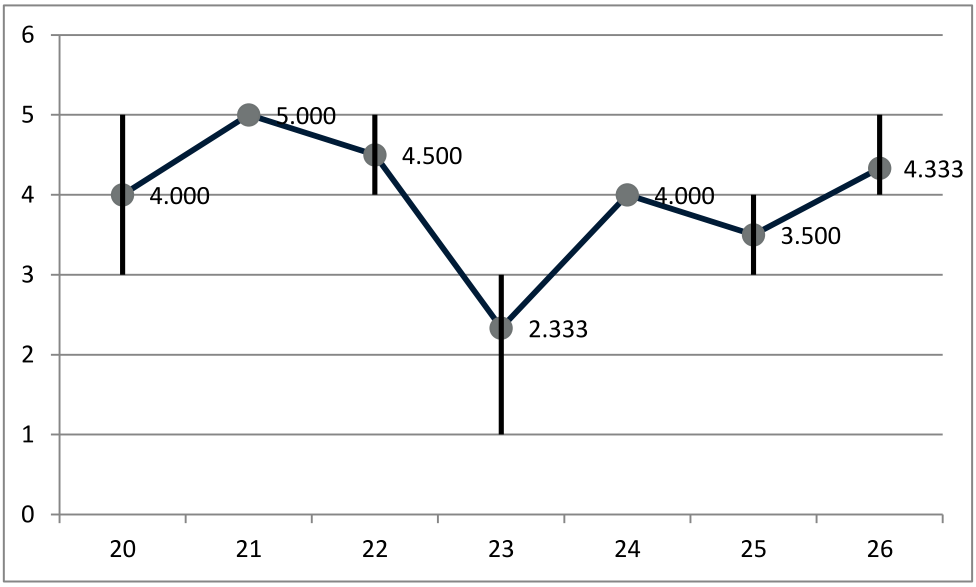

Figure 23. Satisfaction

Figure 23 shows the evaluation of system satisfaction. Except question 23 and question 25, all questions score 4 or above. Question 23 asks if the system works the way wanted. The lowest score given to this question is 1, which is again given by the female evaluator who gives the lowest score to question 11 and 15. Thus, I believe the reason should be also the same. Question 25 is asking about if they want to have the system. The evaluators do not really want to own the system, because they think the system is immature to be released at current stage.

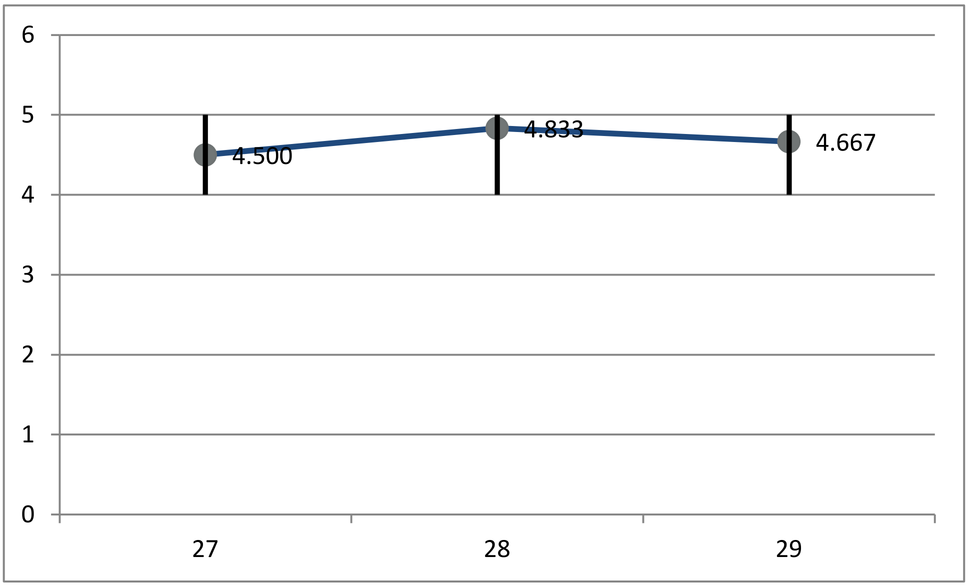

Figure 24. Overall Scenario Evaluation

The overall scenario evaluation is given by Figure 24, and the results indicate all the evaluators are quite satisfied with doing the scenario tasks using the Virtual Fitting Room system. That is, the features of the Virtual Fitting Room system are quite satisfied by general users. Hence, the project objectives are achieved.

8. CONCLUSION

This report presents the observations in clothes shopping experiences and figures out the potential problems. The Virtual Fitting Room system is introduced and implemented which is aimed to increase customers' satisfaction during clothes shopping. The system supports not only basic clothes trying on feature but also functionalities of photo taking and photo downloading using QR code.

The report also presents the system design, the preparation process and the iterative implementation process of the Virtual Fitting Room system. In addition, different types of evaluations have been done, and the final result of the system evaluation shows that the system meets general people's needs, although it is still in an immature stage. The current version of the system does not support 3D clothes modeling and animations. Users feel inflexible with 2D clothes textures. Future works can be done on creating 3D clothes models to further enhance the usability of the system. After all, the Virtual Fitting Room system brings a new way of real-time virtual trying on experience to the customers.

9. REFERENCE

Bhuiyan, M., & Picking, R. (2009). Gesture Control User Interface, what have we done and what's next? *Proceedings of the fifth collaborative research symposium on security, E-learning, Internet and Networking*.

Bhuiyan, M., & Picking, R. (2009). Gesture Controlled User Interfaces for elderly and Disabled. *the Second IASTED International Conference on Telehealth and Assistive Technology*.

Bhuiyan, M., & Picking, R. (2009). Gesture Controlled User Interfaces for Inclusive Design. *12th IFIP Conference on Human Computer Interaction*.

Bhuiyan, M., & Picking, R. (2011). A Gesture Controlled User Interface for Inclusive Design and Evaluative Study of Its Usability. *The International Review on Computers and Software*.

Cheng, C.-M., Chung, M.-F., Yu, M.-Y., Ouhyoung, M., Chu, H.-H., & Chuang, Y.-Y. (2008, 4). ChroMirror: A Real-Time Interactive Mirror for Chromatic and Color-Harmonic Dressing. *CHI '08 extended abstracts on Human factors in computing systems*.

Cho, C.-S., Park, J.-Y., Boeing, A., & Hingston, P. (2010, 8). An implementation of a garment-fitting simulation system using laser scanned 3D body data. *Computers in IndustrY*.

Chu, M., Dalal, B., Walendowski, A., & Begole, B. (2010, 4). Countertop Responsive Mirror: Supporting Physical Retail Shopping for Sellers, Buyers and Companions. *Proceedings of the 28th international conference on Human factors in computing systems*.

Du, H., Henry, P., Ren, X., Cheng, M., Goldman, D. B., Seitz, S. M., & Fox, D. (2011, 9). RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments. *Proceedings of the 13th international conference on Ubiquitous computing*.

Ehara, J., & Saito, H. (2005, 12). Texture Overlay onto Deformable Surface for Virtual Clothing. *Proceedings of the 2005 international conference on Augmented tele-existence*.

Eisert, P., Rurainsky, J., & Fechteler, P. (2007, 9). Virtual Mirror: Real-Time Tracking of Shoes in Augmented Reality Environments. *Image Processing, IEEE International Conference*.

Fuhrmann, A., Groß, C., Luchas, V., & Andreas, W. (2003). Interaction-free dressing of virtual humans. *Computers & Graphics*.

Gavrila, D. M., & Davis, L. S. (1996, 6). 3-D model-based tracking of humans in action: a multi-view approach. *Proceedings of the 1996 Conference on Computer Vision and Pattern Recognition*.

Huang, Y., & Kallmann, M. (2009, 7). Interactive Demonstration of Pointing Gestures fo Virtual Trainers. *Proceedings of the 13th International Conference on Human-Computer Interaction*.

Kallio, S., Kela, J., Jani, M., & Plomp, J. (2006, 5). Visualization of Hand Gestures for Pervasive Computing. *Proceedings of the working conference on Advanced visual interfaces*.

Khoshelham, K. (2011). Accuracy analysis of kinect depth data. *ISPRS workshop laser scanning*.

Kim, D., & Kim, D. (2006, 11). An Intelligent Smart Home Control Using Body Gestures. *Proceedings of the 2006 International Conference on Hybrid Information Technology* .

Lee, M. W., & Nevatia, R. (2007, 2). Body Part Detection for Human Pose Estimation and Tracking. *Proceedings of the IEEE Workshop on Motion and Video Computing*.

Liew, J. S., Kaziunas, E., Liu, J., & Zhuo, S. (2011, 5). Socially-interactive dressing room: an iterative evaluation on interface design. *Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems*.

Lin, L., Wnag, Y., Liu, Y., Xiong, C., & Zeng, K. (2009, 1). Marker-less registration based on template tracking for augmented reality. *Multimedia Tools and Applications*.

Maes, P., Darrel, T., Blumberg, B., & Pentland, A. (1995, 4). The ALIVE System: Full-body Interaction with Autonomous Agents. *Proceedings of the Computer Animation*.

Magnenat-Thalmann, N., Seo, H., & Cordier, F. (2004, 9). Automatic modeling of virtual humans and body clothing. *Journal of Computer Science and Technology*.

Malik, S., & Laszlo, J. (2004, 10). Visual Touchpad: A Two-handed Gestural Input Device. *Proceedings of the 6th international conference on Multimodal interfaces*.

Meng, Y., Mok, P., & Jin, X. (2010, 4). Interactive virtual try-on clothing design systems. *Computer-Aided Design*.

Pavlovic, V. I., Sharma, R., & Huang, T. S. (1997, 7). Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Poppe, R. (2007, 10-11). Vision-based human motion analysis: An overview. *Computer Vision and Image Understanding*, pp. 4-18.

Protopsaltou, D., Luible, C., Arevalo, M., & Magnenat-Thalmann, N. (2002). A body and garment creation method for an Internet based virtual fitting room. *Computer Graphics International 2002 (CGI '02), Springer*, pp. 105-122.

Santangelo, A., Augoello, A., Sorce, S., Pilato, G., Gentile, A., Genco, A., & Gaglio, S. (2007, 11). A virtual shopper customer assistant in pervasive environments. *Proceedings of the 2007 OTM confederated international conference on On the move to meaningful internet systems*.

Shen, E., Lieberman, H., & Lam, F. (2007, 1). What am I gonna wear?: scenario-oriented recommendation. *Proceedings of the 12th international conference on Intelligent user interfaces*.

Tsujita, H., Tsukaka, K., Kambara, K., & Siio, I. (2010, 5). Complete fashion coordinator: a support system for capturing and selecting daily clothes with social networks. *Proceedings of the International Conference on Advanced Visual Interfaces*.

Vera, L., Gimeno, J., Coma, I., & Fernández, M. (2011, 9). Augmented Mirror: Interactive Augmented Reality based on Kinect. *Proceedings of the 13th IFIP TC 13 international conference on Human-computer interaction*.

Volino, P., Cordier, F., & Magnenat-Thalmann, N. (2005, 5). From early virtual garment simulation to interactive fashion design. *Computer-Aided Design*.

Wacker, M., Keckeisen, M., Stoev, S. L., & Straßer, W. (2003, 10). A Comparative Study on User Performance in the Virtual Dressmaker Application. *Proceedings of the ACM symposium on Virtual reality software and technology*.

Wren, C. R., Azarbayejani, A., Trevor, D., & Pentland, A. P. (1997, 7). Pfinder: Real-Time Tracking of the Human Body. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Zhang, W., Matsumoto, T., Liu, J., Chu, M., & Begole, B. (2008, 1). An Intelligent Fitting Room Using Multi-Camera Perception. *Proceedings of the 13th international conference on Intelligent user interfaces*.

Questionnaire For Conceptual Model

		1	2	3	4	5	
Overall, how do you find the user interface							
	Terrible	☐	☐	☐	☐	☐	Wonderful
	Difficult	☐	☐	☐	☐	☐	Easy
	Frustrating	☐	☐	☐	☐	☐	satisfying
	Inadequate power	☐	☐	☐	☐	☐	Adequate power
	Dull	☐	☐	☐	☐	☐	Interesting
	Rigid	☐	☐	☐	☐	☐	Flexible
Screen							
Organization of information	Confusing	☐	☐	☐	☐	☐	Clear
Sequence of screens	Confusing	☐	☐	☐	☐	☐	Clear
Learning							
Learning to operate the user interface	Difficult	☐	☐	☐	☐	☐	Very Easy
Exploring new features by trial and error	Difficult	☐	☐	☐	☐	☐	Very Easy
Performing tasks is straightforward	Never	☐	☐	☐	☐	☐	Always

Which functionalities of the interface did you try?		
☐ Color buttons	☐ Up arrow	☐ Down arrow
☐ Photo buttons	☐ QR code	
How many colors did you tried when you play with the user interface?		
If all, why did you try them all?		
What are the colors that you tried and what are the characters associated with those colors that you can remember?		
List the most NEGATIVE aspects		
1. 2.		
List the most POSITIVE aspects		
1. 2.		

-------------------------- THE END --------------------------

APPENDIX II

Before-Scenario questionnaire

How often do you go shopping? (Choose ONE)

a) Everyday
b) Once per week
c) Once per month
d) Less that once per month

When did you go shopping?(Choose ONE or MORE)

- ☐ Weekend
- ☐ Public holiday
- ☐ Friday afternoon
- ☐ Holiday eve
- ☐ Weekday
- ☐ Other: ____

On each shopping day, how many clothes did you try on average? (Choose ONE)

a) Less than 5
b) 5 to 10
c) 11 to 20
d) More than 20

On each shopping day, how much did you spend on average? (Choose ONE)

a) Less than 100
b) 100-500
c) 500-1000
d) More than 1000

How did you make purchasing decisions? (Choose ONE or MORE)

- ☐ Try the clothes on and decide by myself
- ☐ Ask friends who go shopping together
- ☐ Call someone for opinion
- ☐ Send message to someone for opinion
- ☐ Take a photo and show/send it to someone for opinion
- ☐ Other: ____

List two of the most time consuming activities during clothes shopping?

1.____

2. ____

Do you know what QR code is?

Do you have any QR code scanner app in you cell phone?

------------------------------ THE END ------------------------------

APPENDIX III

Instructions for Task Centered Walkthrough

Imagine that you are in a clothing shop with a new system called "Virtual Fitting Room". Without using the real fitting room, you are able to try on clothes virtually. Also you can take photos of your virtual appearance by using it. Now you are asked to accomplish several tasks using the system.

Scenario Task 1:
Let the system recognize you so that you can control and play with it.

Scenario Task 2:
Try on some dresses.

Scenario Task 3:
Try more dresses by using the navigation arrows.

Scenario Task 4:
Take a photo.

Scenario Task 5:
Download and save your photo using your cell phone.

Scenario Task 6:
Send your photo to a friend.

Scenario Task 7:
Give up the system control by just walking away.

NOTE: Your photos will be removed immediately after you finish this scenario, so that you do not need to worry out the privacy issues.

-------------------------------- THE END --------------------------------

APPENDIX IV

After Scenario Questionnaire

Usefulness		1	2	3	4	5	
1. It helps me be more effective.	Disagree	☐	☐	☐	☐	☐	Agree
2. It is useful	Disagree	☐	☐	☐	☐	☐	Agree
3. It makes the tasks I want to accomplish easier to get done.	Disagree	☐	☐	☐	☐	☐	Agree
4. It saves me time when I use it.	Disagree	☐	☐	☐	☐	☐	Agree
5. It meets my needs.	Disagree	☐	☐	☐	☐	☐	Agree
6. It does everything I would expect it to do.	Disagree	☐	☐	☐	☐	☐	Agree
Ease of Use		**1**	**2**	**3**	**4**	**5**	
7. It is easy to use.	Disagree	☐	☐	☐	☐	☐	Agree
8. It is simple to use.	Disagree	☐	☐	☐	☐	☐	Agree
9. It is user friendly.	Disagree	☐	☐	☐	☐	☐	Agree
10. It requires the fewest steps possible to accomplish what I want to do with it.	Disagree	☐	☐	☐	☐	☐	Agree
11. It is flexible.	Disagree	☐	☐	☐	☐	☐	Agree
12. Using it is effortless.	Disagree	☐	☐	☐	☐	☐	Agree
13. I can use it without written instructions.	Disagree	☐	☐	☐	☐	☐	Agree
14. I don't notice any inconsistencies as I use it.	Disagree	☐	☐	☐	☐	☐	Agree
15. Both occasional and regular users would like it.	Disagree	☐	☐	☐	☐	☐	Agree
16. I can use it successfully every time.	Disagree	☐	☐	☐	☐	☐	Agree
Ease of Learning		**1**	**2**	**3**	**4**	**5**	
17. I learned to use it quickly.	Disagree	☐	☐	☐	☐	☐	Agree
18. I easily remember how to use it.	Disagree	☐	☐	☐	☐	☐	Agree
19. It is easy to learn to use it.	Disagree	☐	☐	☐	☐	☐	Agree
Satisfaction		**1**	**2**	**3**	**4**	**5**	
20. I am satisfied with it.	Disagree	☐	☐	☐	☐	☐	Agree
21. I would recommend it to a friend.	Disagree	☐	☐	☐	☐	☐	Agree
22. It is fun to use.	Disagree	☐	☐	☐	☐	☐	Agree
23. It works the way I want it to work.	Disagree	☐	☐	☐	☐	☐	Agree
24. It is wonderful.	Disagree	☐	☐	☐	☐	☐	Agree
25. I feel I need to have it.	Disagree	☐	☐	☐	☐	☐	Agree
26. It is pleasant to use.	Disagree	☐	☐	☐	☐	☐	Agree
Overall		**1**	**2**	**3**	**4**	**5**	
27. I am satisfied with the ease of completing the tasks in this scenario.	Disagree	☐	☐	☐	☐	☐	Agree
28. I am satisfied with the amount of time it took to complete the tasks in this scenario.	Disagree	☐	☐	☐	☐	☐	Agree
29. I am satisfied with the support information (helping message, instructions) when completing the tasks	Disagree	☐	☐	☐	☐	☐	Agree

---------------------------------- THE END ----------------------------------